

Survey Research

“A comprehensive system for collecting information to describe, compare or explain knowledge, attitudes and behaviour over large populations”

- good for
 - Investigating the nature of a large population
 - Testing theories where there is little control over the variables
- limitations
 - Relies on self-reported observations
 - Difficulties of sampling and self-selection
 - Information collected tends to subjective opinion

See:

Shari Lawrence Pfleeger and Barbara A. Kitchenham, "Principles of Survey Research," Software Engineering Notes, (6 parts) Nov 2001 - Mar 2003

What is Survey Research?

- Survey Research ≠ Questionnaires
 - Can use questionnaires in any method
 - E.g. pre- and post- test in experiments
 - Can do survey research without questionnaires
 - E.g. using interviews, data logging, etc
- Distinguishing features:
 - Precondition: a clear research question that asks about the nature of a particular target population
 - selection of a representative sample from a well-defined population
 - data analysis techniques used to generalize from that sample to the population
 - Most suitable for answering base-rate questions

When to use Survey Research

- To evaluate the **frequency** of some characteristic across a population
 - E.g. how many companies use agile methods?
- To evaluate the **severity** of some condition that occurs in a population
 - E.g. what's the average cost overrun of software projects?
- To identify factors that **influence** a characteristic or condition
 - E.g. What factors cause companies to adopt new ASE tools?

Starting point

- Set clear objectives
 - A hypothesis to be tested
 - Any alternative explanations to be investigated
 - Identify a scope for the study appropriate for the objectives
 - Identify resources needed to meet the objectives
- Check that a survey is the right method:
 - Is it clear what population can answer the questions reliably?
 - Is there a way to get a representative sample of that population?
 - Do you have resources to obtain a large enough sample?
 - Is it clear what variables need to be measured?
 - Is it clear how to measure them?

I. Probability Sampling

- Simple random sampling
- Stratified random sampling
- Systematic random sampling
- Cluster random sampling
- Multi-stage sampling

5

II. Non-probability Sampling

- Convenience sampling
- Purposive sampling
- Expert sampling
- Quota sampling
- Modal sampling
- Heterogeneity sampling
- Snowball sampling

6

Avoiding Sampling Bias

- Clear definition of the survey sample:
 - Define the U, the unit of analysis
 - Define the P, the target population
 - ...such that $P = \{U\}$
 - Sample of the entire target population
 - not just the most accessible portion of it!
- Stratified Random Sampling for confounding variables:
 - E.g. U = individual developer,
P = developers working in Canadian software companies
 - ... but what if 80% of our sample comes from a single, dominant company?
 - If we really wanted U = Canadian Software Companies
 - Then change P
 - Alternatively, if company is a confounding variable
 - Group population by company, then sample within each

Survey Study Designs

- Cross-sectional design
 - Used to obtain a snapshot of participants' current activities.
- Case-control design
 - Asks each participant about several related issues
 - Used to establish whether a correlation exists between certain phenomena, across the population.
- Longitudinal study
 - Administer a survey periodically to track changes over time
- Cohort study
 - A longitudinal study that tracks the same participants each time

Avoiding Self-selection Bias

- Sampling the right population might not be enough
 - Low response rates (e.g. under 10%) are common
 - Low response rates may invalidate the sampling method
 - Participants who choose to respond might be unrepresentative:
 - E.g. People who are least busy
 - E.g. People who have a strong opinion on the research topic
- Probe reasons for low response rate
 - E.g. follow up phone calls to non-respondents

Create a survey instrument

- Use/adapt other people's instruments if possible
 - Existing instruments have already been validated
 - Makes it easier to compare research results
- Challenges:
 - Phrase the questions so all participants understand them in the same way
 - Closed questions:
 - Hard to give appropriate choices of answer
 - Hard to ensure all respondents understand the choices in the same way
 - Open questions:
 - Hard to analyse the responses
- Prototyping and validation
 - Test that participants can understand the questions
 - Test how long it takes them to answer
 - Use prototyping results to improve the instrument

Question Design

- Questions must be unambiguous and understandable:
 - Language appropriate to the population
 - Use standard grammar, punctuation, spelling
 - Each question covers exactly one concept
 - Avoid vague or ambiguous qualifiers
 - Balance positive and negative questions
- Typical mistakes:
 - Questions that participants can't answer
 - E.g. asking about decisions they weren't involved in
 - Double edged questions
 - E.g. "have you used SE tools or techniques, or would you consider using them?"
 - Leading questions
 - E.g. "did the project fail because of poorly managed requirements?"
 - Appropriation - reinterpreting participants' responses

Answer Design

- Response Categories
 - Numeric values (e.g. number of months on the project)
 - Nominal categories (e.g. type of software being built)
 - Binary (e.g. Yes/No)
 - Ordinal scales (e.g. "how strongly do you agree with this statement...")
- Response options should be:
 - Exhaustive (but not too long!)
 - Include 'other' if you cannot ensure they are exhaustive
 - Mutually exclusive
 - Allow for multiple selections if appropriate
- Using ordinal scales:
 - Use 5 - 7 points on the scale
 - Label the points on the scale with words
 - End points must mean the opposite of one another
 - Intervals must seem to be evenly spaced

Reliability

- Test-Retest Reliability
 - If the same person answers the survey twice, do you get the same answers?
 - Problems:
 - What if the world has changed?
 - What if answering the questionnaire changes their attitude?
 - What if they just remember their answers from last time?
- Alternate Form Reliability
 - Do re-worded or re-ordered questions yield the same results?
- Inter-rater Reliability
 - If someone else administers the questions, do you get the same answers?
 - If someone else codes the responses, do you get the same results?

Surveys vs. other methods

- Use survey research when:
 - You need to find out what's true across (some part of) the s/w industry
 - Establish what is normal, common or uncommon.
- Use case study when:
 - You want to understand what developers actually do
 - deeper insights into what happens in a small number of selected cases.
- Use an experiment (or quasi-experiment) when:
 - You want to investigate whether a particular technique has an effect on quality, development time, etc
 - tests for a causal relationship.
- Use an ethnography when:
 - You want to understand the culture and perspective of developers
 - Probes how developers themselves make sense of their context
- Use action research when:
 - You need to solve a pressing problem, and understand whether the solution was effective
 - Focusses on effecting change, and learning from the experience